
 Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

 (TM)

 Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

 Msearch-pp protein - protein database search, using Smith-Waterman algorithm
 Run on: Sat May 13 10:16:40 2000; MasPar time 4.56 Seconds
 Tabular output not generated. 327.460 Million cell updates/sec

Title: >US-09-331-631-3
 Description: (186-248) from US09331631.pep (5 of 5)
 Perfect Score: 492
 Sequence: 1 KRDPQOREYEDCDRRRCEQOE.....LINPORGSGRYEGEGEKOS 63

Scoring table: PAM 150
 Gap 11

Searched: 188963 seqs, 23686106 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: a-geneseq35
 1:geneseqp

Statistics: Mean 25.162; Variance 101.062; scale 0.249

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description	Pred. No.
1	492	100.0	666	1	Macadamia integrifolia	6.47e-40
2	462	93.9	625	1	Macadamia integrifolia	8.17e-37
3	449	91.3	666	1	Macadamia integrifolia	1.79e-35
4	178	36.2	525	1	Theobroma cacao anti	2.98e-08
5	178	36.2	566	1	Sequence encoded by 67	2.98e-08
6	134	31.3	590	1	Gossypium hirsutum ant	5.58e-06
7	135	27.4	637	1	Hordeum vulgare anti	3.23e-04
8	103	20.9	593	1	Zea mays antimicrobial	2.37e-01
9	92	18.7	918	1	Human androgen recepto	2.08e+00
10	92	18.7	919	1	Androgen receptor.	2.08e+00
11	92	18.7	919	1	Human androgen recepto	2.08e+00
12	86	17.5	112	1	Y04866	6.60e+00
13	86	17.5	126	1	Mycobacterium species	6.60e+00
14	85	17.3	2289	1	Protein derived from r	7.99e+00
15	84	17.1	33	1	Zea mays antimicrobial	9.66e+00
16	84	17.1	33	1	Antimicrobial maize pe	9.66e+00
17	84	17.1	919	1	Human androgen recepto	9.66e+00
18	83	16.9	78	1	Stenocarpus sinuatus a	1.17e+01
19	83	16.9	816	1	Spinocardiellat ataxia	1.17e+01
20	82	16.7	516	1	Soybean glycinin A3B4	1.41e+01
21	82	16.7	898	1	Mycobacterium tubercul	1.41e+01
22	81	16.5	669	1	Mouse liver cancer-ori	1.70e+01
23	80	16.3	205	1	Human secreted protein	2.05e+01

ID	Score	Query Match	Length DB	ID	Description	Pred. No.
24	80	16.3	537	1	R75188	2.05e+01
25	79	16.1	86	1	GST-HD fusion protein	2.47e+01
26	79	16.1	86	1	GST-HD fusion protein	2.47e+01
27	79	16.1	263	1	Moloney murine leukemia	2.47e+01
28	79	16.1	657	1	EMR-1 gene product.	2.47e+01
29	79	16.1	712	1	R29580	2.47e+01
30	78	15.9	55	1	W30749	2.98e+01
31	78	15.9	55	1	W04825	2.98e+01
32	78	15.9	138	1	Murine vascular endoth	2.98e+01
33	78	15.9	140	1	W84159	2.98e+01
34	78	15.9	145	1	Y04885	2.98e+01
35	78	15.9	147	1	W86212	2.98e+01
36	78	15.9	150	1	W86211	2.98e+01
37	78	15.9	152	1	W86210	2.98e+01
38	78	15.9	155	1	W86209	2.98e+01
39	78	15.9	160	1	W86208	2.98e+01
40	78	15.9	167	1	W86234	2.98e+01
41	78	15.9	188	1	W04826	2.98e+01
42	78	15.9	188	1	W86201	2.98e+01
43	78	15.9	188	1	W80490	2.98e+01
44	78	15.9	195	1	W80491	2.98e+01
45	78	15.9	539	1	W33628	2.98e+01

ALIGNMENTS

RESULT 1
 ID W62829 standard; Protein: 666 AA.
 AC W62829:
 DT 27-OCT-1998 (first entry)
 DE Macadamia integrifolia antimicrobial protein.
 KW antimicrobial protein; infestation: control.
 OS Macadamia integrifolia.
 FH Key Location/Qualifiers
 FT Peptide 1..28
 FT Protein /note="signal peptide"
 FT /note="mature protein"
 PN W09827805-A1.
 PD 02-JUL-1998.
 PF 22-DEC-1997: AU0874.
 PR 20-DEC-1996: AU-004275.
 PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
 PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
 DR WPI: 98-372729/32.
 DR N-PSDB: V42311.
 PT Novel anti-microbial protein from e.g. Macadamia integrifolia -
 useful for controlling microbial infestations of plants or mammals
 PS Claim 1; Page 39-41; 96pp; English.
 CC The sequence is that of an antimicrobial protein which can
 CC be used to control microbial infestations in plants and mammalian
 CC animals.
 SO Sequence 666 AA:
 Query Match 100.0%; Score 492; DB 1; Length 666;
 Best Local Similarity 100.0%; Pred. No. 6.47e-40;
 Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 186 KRDPQOREYEDCDRRRCEQOEPRQOYOCORCRQOHORGDLINPORGSGRYEGEGE 245
 QY 186 KRDPQOREYEDCDRRRCEQOEPRQOYOCORCRQOHORGDLINPORGSGRYEGEGE 245
 DB 246 KOS 248
 QY 246 KOS 248
 RESULT 2
 ID W62830 standard; Protein: 625 AA.
 AC W62830:
 DT 27-OCT-1998 (first entry)
 DE Macadamia integrifolia antimicrobial protein.
 KW antimicrobial protein; infestation: control.

OS Macadamia integrifolia.
FH Key Location/Qualifiers
FT Peptide 1..28
FT /note= "signal peptide"
FT Protein 29..666
FT /note= "mature protein"
PN W09827805-A1.
PD 02-JUL-1998.
PF 22-DEC-1997; AU0874.
PR 20-DEC-1996; AU-004275.
PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
DR WPI: 98-377279/32.
DR N-PSDB: V42316.
PT Novel anti-microbial protein from e.g. Macadamia integrifolia -
PT useful for controlling microbial infestations of plants or mammals
PS Claim 1; Page 43-45; 96pp; English.
CC The sequence is that of an antimicrobial protein which can
CC be used to control microbial infestations in plants and mammalian
CC animals.
SQ Sequence 625 AA;

Query Match 93.9%; Score 462; DB 1; Length 625;
Best Local Similarity 93.7%; Pred. No. 8.17e-37;
Matches 59; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

DB 145 KRDPQREYEDCRRCCEQEPRLQYQOCRCCEQORHGSGDLNMPORGSGRYEEGEE 204
|||
QY 166 KRDPQREYEDCRRCCEQEPRLQYQOCRCCEQORHGSGDLNMPORGSGRYEEGEE 245

DB 205 KOS 207
|||
QY 246 KOS 248

RESULT 3
ID W62828 standard; Protein; 666 AA.
AC W62828;
DT 27-OCT-1998 (first entry)
DE Macadamia integrifolia antimicrobial protein.
KW antimicrobial protein; infestation; control.
OS Macadamia integrifolia.
FH Key Location/Qualifiers
FT Peptide 1..28
FT /note= "signal peptide"
FT Protein 29..666
FT /note= "mature protein"
PN W09827805-A1.
PD 02-JUL-1998.
PF 22-DEC-1997; AU0874.
PR 20-DEC-1996; AU-004275.
PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
DR WPI: 98-377279/32.
DR N-PSDB: V42310.
PT Novel anti-microbial protein from e.g. Macadamia integrifolia -
PT useful for controlling microbial infestations of plants or mammals
PS Claim 1; Page 34-36; 96pp; English.
CC The sequence is that of an antimicrobial protein which can
CC be used to control microbial infestations in plants and mammalian
CC animals.
SQ Sequence 666 AA;

Query Match 91.3%; Score 449; DB 1; Length 666;
Best Local Similarity 92.1%; Pred. No. 1.79e-35;
Matches 58; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

DB 186 KRDPQREYEDCRRCCEQEPRLQYQOCRCCEQORHGSGDLNMPORGSGRYEEGEE 245
|||
QY 186 KRDPQREYEDCRRCCEQEPRLQYQOCRCCEQORHGSGDLNMPORGSGRYEEGEE 245

DB 246 EQS 248
|||

QY 246 KOS 248

RESULT 4
ID W62831 standard; Protein; 525 AA.
AC W62831;
DT 27-OCT-1998 (first entry)
DE Theobroma cacao antimicrobial protein.
KW antimicrobial protein; infestation; control.
OS Theobroma cacao.
PN W09827805-A1.
PD 02-JUL-1998.
PF 22-DEC-1997; AU0874.
PR 20-DEC-1996; AU-004275.
PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
DR WPI: 98-377279/32.
PT Novel anti-microbial protein from e.g. Macadamia integrifolia -
PT useful for controlling microbial infestations of plants or mammals
PS Claim 1; Page 47-49; 96pp; English.
CC The sequence is that of an antimicrobial protein which can
CC be used to control microbial infestations in plants and mammalian
CC animals.
SQ Sequence 525 AA;

Query Match 36.2%; Score 178; DB 1; Length 525;
Best Local Similarity 37.7%; Pred. No. 2.98e-08;
Matches 23; Conservative 20; Mismatches 15; Indels 3; Gaps 3;

DB 82 ORQYQCCGRCQEQOQOCRCCEQORHGSGDLNMPORGSGRYEEGEE 141
|||
QY 191 QREYEDCRRC-EOQE-PROQYQOCRCCEQORHGSGDLNMPORGSGRYEEGEE 247

DB 142 N 142
|||
QY 248 S 248

RESULT 5
ID R20181 standard; Protein; 666 AA.
AC R20181;
DT 16-APR-1992 (first entry)
DE Sequence encoded by 67 kD T. cacao protein cDNA.
KW Cocoa; flavour; vicillin; seed storage protein.
OS Theobroma cacao.
PN W09119801-A.
PD 26-DEC-1991.
PF 07-JUN-1991; G00914.
PR 11-JUN-1990; GB-013016.
PA (MNSC) MARS UK LTD.
PI Spencer ME, Hodge R, Deakin EA, Ashton S;
DR WPI: 92-024418/03.
DR N-PSDB: 020377.
PT Recombinant cocoa proteins - are responsible for flavour in cocoa
PT beans and produced in large quantities using yeast and bacterial
PT expression vectors
PS Claim 4; Fig 2; 59pp; English.
CC The inventors claim a 67 kD and 31 kD T. cacao protein, and
CC fragments, and encoding DNAs. The 47 kD and 31 kD proteins are
CC derived from the 67 kD precursor. T. cacao protein cDNA was
CC detected in a cDNA library prepared from immature cocoa beans RNA
CC using a probe based on the AA sequence of a CNBR peptide common to
CC the 47 kD and 31 kD polypeptides. Homology searches revealed close
CC homologues between the 67 kD polypeptide and the vicillins, which are
CC seed storage proteins.
SQ Sequence 566 AA;

Query Match 36.2%; Score 178; DB 1; Length 566;
Best Local Similarity 37.7%; Pred. No. 2.98e-08;
Matches 23; Conservative 20; Mismatches 15; Indels 3; Gaps 3;

DB 82 ORQYQCCGRCQEQOQOCRCCEQORHGSGDLNMPORGSGRYEEGEE 141
|||
QY 191 QREYEDCRRC-EOQE-PROQYQOCRCCEQORHGSGDLNMPORGSGRYEEGEE 247

QY	191	QREYDCRRRC-EGGE-PRQYQCORRCERQNRHGRCG-DLIMPGRSGSRYEEGEKQ	247
Db	142	N 142	
QY	248	S 248	

RESULT 6
ID M62832 standard: Protein; 590 AA.
AC M62832:
DE 27-OCT-1998 (first entry)
DE Gossypium hirsutum antimicrobial protein.
KW antimicrobial protein; infestation; control.
OS Gossypium hirsutum.
PN M09827805-A1.
PD 02-JUL-1998.
PF 22-DEC-1997: AU0874.
PR 20-DEC-1996: AU-004275
RA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
DR WPI: 98-37729/932.
PT Novel anti-microbial protein from e.g. Macadamia integrifolia -
PT useful for controlling microbial infestations of plants or mammals
PS Claim 1: Page 49-51; 36pp: English.
CC The sequence is that of an antimicrobial protein which can
CC be used to control microbial infestations in plants and mammalian
CC animals.
SQ Sequence 590 AA;

RESULT 7
ID M62837 standard: Protein; 637 AA.
AC M62837;
DT 27-OCT-1998 (first entry)
DE Hordeum vulgare antimicrobial protein.
KW antimicrobial protein; infestation; control.
OS Hordeum vulgare.
PN M09827805-A1.
PD 02-JUL-1998.
PE 22-DEC-1997;
PR 20-DEC-1996; AU-004275.
RA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOG.
PI Bover NI, Goulier KC, Green JL, Manners JM, Marcus JP;
DR WPI: 98-317219/32.
PT Novel anti-microbial protein from e.g. *Macadamia integrifolia* -
PT useful for controlling microbial infestations of plants or mammals
PS Claim 1: Page 60-62; 96pp: English.
CC The sequence is that of an antimicrobial protein which can
CC be used to control microbial infestations in plants and mammalian
CC animals. Sequence 637 AA;
50

RESULT	8
ID	W62835 standard; Protein; 593 AA.
AC	W62835;
DT	27-OCT-1998 (first entry)

DE zea mays antimicrobial protein.
KW antimicrobial protein; infestation; control.
OS zea mays.
PS WO9827805-A1.
PN 02-JUL-1998.
PD 22-DEC-1997; AU0874.
PF 20-DEC-1996; AU-004275.
PR (BETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
PA BOWER NI, GOULTER KC, GREEN JL, MANNERS JM, MARCUS JP;
PI WPI; 98-377219/32.
DR Novel anti-microbial protein from e.g. *Macadamia integrifolia* -
PT useful for controlling microbial infestations of plants or mammals
PS Claim 1; Page 58-60; 96pp; English.
CC The sequence is that of an antimicrobial protein which can
CC be used to control microbial infestations in plants and mammalian
CC animals.
SQ Sequence 593 AA;

RESULT	9	standard: Protein; 918 AA.
ID	R12223	
AC	R12223;	
DT	20-AUG-1991	(first entry)
DE	Human androgen receptor.	
KW	hAR; DNA-binding protein; steroid hormone.	
OS	Homo sapiens.	
FT	key	Location/Qualifiers
FT	domain	556..626
FT		/label= DNA-binding domain
FT		/note= "cysteine-rich"
FN	WO9107423-A.	
PD	30-MAY-1991.	
PF	19-OCT-1990.	006015.
PR	17-NOV-1989:	US-438775.
PA	(ARCH-) ARCH DEV CORP.	
P1	Liao S, Chang C;	
DR	wpi; 91-178048/24.	
DR	N-PSDE: Q12001.	
PT	Androgen receptor and TR2 DNA binding proteins - DNA sequences	
PT	and antibodies for detection and quantification methods	
PS	Claim 25; Fig 3; 79pp; English.	
CC	This sequence was deduced from a cDNA clone isolated by screening	
CC	commercially available human testis and prostate lambda gtl1 cDNA	
CC	libraries. The sequence is very similar to that of rat AR and in	
CC	CC the DNA-binding domain it is identical to that of rat DNA-binding	
CC	domain. Homology comparisons with other known steroid receptors	
CC	indicate that hAR is more closely related to glucocorticoid,	
CC	mineralo-corticoid and progesterone receptors than to v-erb-A or to	
CC	receptors for oestrogen, vitamin D and thyroid hormones.	
Q0	Sequence 918 AA;	

RESULT	10
ID	W14783 standard; Protein; 919 AA
AC	W14783; [†]
DI	22-JUN-1997 (first entry)
DE	Androgen receptor.

RESULT 14

ID W14987 standard; Protein; 2289 AA.
AC W14987:
DT 02-FEB-1998 (first entry)
DE Protein derived from regulatory gene from U. maydis.
KW Ustilago maydi; maize; pathogen; fungus; fungicide; mutant.
OS Ustilago maydis.
PN W09712911-A1.
PD 10-APR-1997.
PF 30-SEP-1996; E04254.
PR 25-MAR-1996; DE-011758.
PR 04-OCT-1995; DE-036890.
PA (BADI) BASF AG.
PI Rahmann R, Quadbeck-seeger C;
DR WPI: 97-226165/20.
DR N-PSDB; T62940.
PT New regulatory gene from Ustilago maydis and derived protein -
PT useful as target for identifying fungicides, also mutant fungus for
PT infecting maize
PS Claim 9; Page 9-18; 42pp; German.
CC The gene is a regulatory gene in U. maydis (a maize pathogen)
CC and probably has a repressive action.
CC Ustilago strains with a mutation in this gene can form galls, but not
CC spores, in maize, so the galls (used as food) are visibly detectable.
SQ Sequence 2289 AA;

Query Match 17.3%; Score 85; DB 1; Length 2289;
Best Local Similarity 37.0%; Pred. No. 7.99e+00;

Matches 10; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

Db 2027 RDERERERDGLDARDGRGRYGD 2053
QY 217 REQORHGRGDLNPRGSGRREGE 243

RESULT 15

ID W62836 standard; Protein; 33 AA.
AC W62836:
DT 27-OCT-1998 (first entry)
DE Zea mays antimicrobial protein.
KW antimicrobial protein; infestation; control.
OS Zea mays.
PN W09827805-A1.
PD 02-JUL-1998.
PF 22-DEC-1997; AU0874.
PR 20-DEC-1996; AU-004275.
PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
DR WPI: 98-377279/32.
PT Novel anti-microbial protein from e.g. Macadamia integrifolia -
PT useful for controlling microbial infestations of plants or mammals
PS Disclosure; Page 60; 96pp; English.
CC The sequence is that of an antimicrobial protein which can
CC be used to control microbial infestations in plants and mammalian
CC animals.
SQ Sequence 33 AA;

Query Match 17.1%; Score 84; DB 1; Length 33;
Best Local Similarity 44.0%; Pred. No. 9.66e+00;

Matches 11; Conservative 7; Mismatches 4; Indels 3; Gaps 2;

Db 6 ECRROCLRHGQPMWETQECMRRCR 30
QY 196 DCRRC-EGQE--PROGYCCRCR 217

Search completed: Sat May 13 10:16:48 2000
Job time : 8 secs.

